

Automated Application of Clinical Practice Guidelines for Asthma Management

Alan R. Ertle, M.D., Emily M. Campbell, R.N., M.S., William R. Hersh, M.D.
Biomedical Information Communication Center
Oregon Health Sciences University
Portland, Oregon

Clinical practice guidelines will be more beneficial when they are integrated with the electronic medical record. We applied natural language processing (NLP) techniques to extract clinical findings from outpatient progress notes in an attempt to: 1) select outpatient chart notes appropriate for evaluation against the National Heart, Lung, and Blood/National Asthma Education Program clinical practice guidelines for the diagnosis and management of asthma, 2) determine patient need for inhaled anti-inflammatory agents, and 3) quantify the severity of asthma. Our results were compared to judgements by an expert panel of practitioners. We were able to correctly identify the need for inhaled anti-inflammatory agents 76% of the time. The success of this pilot project could have broad implications for the application of clinical practice guidelines.

INTRODUCTION

Clinical practice guidelines are designed to provide clinicians with evidence-based, up-to-date information to promote high quality and cost-effective care.¹ A limitation of current guidelines is that they exist on paper or in computer reproduction of paper forms making interactive use difficult. The major benefit of guidelines, whether used to guide care in real-time or critique it afterwards, will only be achieved when they are integrated into the electronic medical record (EMR). The major goal of this study was to develop and evaluate a tool for predicting severity of asthma based on practice guidelines from clinician outpatient progress notes. Such a tool, if successfully constructed, could provide feedback to clinicians caring for patients with asthma.

Asthma is a highly prevalent disease in the United States. Current estimates indicate approximately 5% of the population has active asthma.^{2,3} Asthma is also expensive. Estimates of asthma cost for 1992 were \$6.2 billion or 1% of U. S. health care costs.⁴ A major change in understanding and focus of treatment of asthma has occurred over the past decade. Asthma is an inflammatory process of the airways and inhaled

anti-inflammatory agents should be an integral part of therapy for many asthmatic patients. Recommendations from the National Heart, Lung, and Blood Institute (NHLBI) in their National Asthma Education Program (NAEP) and in the International Consensus Report emphasize the use of inhaled anti-inflammatory agents for all but mild asthma.⁵ There is strong evidence that there is broad awareness of these guidelines by primary care physicians, but that the acceptance and compliance with the recommendations is significantly less.⁶ Careful application and evaluation of the NHLBI/NAEP guidelines could increase the appropriate use of inhaled anti-inflammatory agents reducing morbidity, mortality, and costs associated with asthma.

There has been significant research to determine the severity of asthma and severity of asthma rating scales.⁷⁻¹⁰ The focus of these asthma rating scales has been on more acute, more severe asthma, and the primary function has been to determine need for hospitalization or subsequent need for additional resources. These severity indexes come in two forms: 1) patient or parent questionnaires based on historical facts and symptoms, and 2) health care provider scales based on physical signs and objective tests. These asthma severity rating scales have the advantage of completeness and codification of the entries. The definitions of asthma severity defined in the NHLBI/NAEP are directed toward the outpatient management of all asthmatics, not just asthmatics who are severe or are having an acute exacerbation.⁵ The information required to define asthma severity by the NHLBI/NAEP guidelines includes frequency of symptoms, activity tolerance, nocturnal symptoms, school/work attendance, and lung function measurements (TABLE 1).

Several computer-based projects have been undertaken to assess and/or improve the quality of care in asthmatic patients. There has been at least one expert system developed to evaluate asthma severity.¹¹ Others have attempted to extract the findings in discharge summaries for entry in a database¹² or partition outpatient records into exacerbation/non-exacerbation

for later manual quality measurement.¹³ One problem with the latter study was reliance on ICD-9 codes to identify patient records related to asthma. This method of record identification is inherently limited by problems with coding such as using these codes for patients with chronic obstructive pulmonary disease or not being able to include all diagnostic codes for a particular outpatient encounter.

TABLE 1.

**NHLBI/NAEP International Consensus Report
Asthma Diagnosis Guidelines**

Mild Asthma

Symptoms: Intermittent, brief (<1 hour)
wheeze/cough/dyspnea up to 2 times weekly
Asymptomatic between exacerbations
Activity Tolerance: Brief (<½ hour)
wheeze/cough/dyspnea with activity
Nocturnal Symptoms: (<2 times a month) nocturnal
cough/wheeze
School/Work: Good school or work attendance
Lung Function (FEV₁ or PEFR): ≥80% patient baseline and
varies less than 20% daily when symptomatic.

Moderate Asthma

Symptoms: More than 1-2 times weekly. Exacerbations affect
sleep or activity level (not asymptomatic between) and
may last several days. Occasional emergency care.
Activity Tolerance: Mild to moderate wheeze/cough/dyspnea with
activity
Nocturnal Symptoms: (More than 2 times a month) nocturnal
cough/wheeze
School/Work: Some affect on school or work attendance
Lung Function (FEV₁ or PEFR): 60-80% patient baseline and
varies 20- 30% daily when symptomatic.

Severe Asthma

Symptoms: Wheeze and dyspnea continuous and daily. Frequent
exacerbations. Occasional hospitalizations and
emergency treatment.
Activity Tolerance: Significantly limited activity level.
Nocturnal Symptoms: Almost nightly nocturnal cough/wheeze
School/Work: Poor school or work attendance
Lung Function (FEV₁ or PEFR): <60% patient baseline and varies
more than 30% daily when symptomatic

It is not clear how much information relevant to and necessary for this asthma guideline will be documented by providers and thus be available in outpatient notes. Documentation of historical facts, symptom, signs and tests can be highly variable. However, if enough information is available in some proportion of outpatient notes for natural language assessment of asthma severity and appropriateness of therapy, then application of these guidelines to many records and patients is possible. Application of this scoring process could be useful in clinical research, health care provider reminders, quality assurance, and scorecards for managed care organizations.

This primary goal of this study was to determine if an automated system could be developed to extract clinical information from textual outpatient notes in order to apply the NHLBI/NAEP guidelines for the diagnosis and management of asthma. Ideally, this methodology would match expert opinion in prediction of need for inhaled anti-inflammatory agents. The first step to accomplish this goal was to construct a method for the identification of appropriate outpatient notes. The second step was the development of an automatic scoring algorithm to determine the need for inhaled anti-inflammatory agents. The "gold standard" used to determine need for inhaled anti-inflammatory agents was a panel of expert pulmonary practitioners. The primary outcome measure was the percent of documents in the test set in which the need for inhaled anti-inflammatory agents was correctly determined by the computer scoring algorithm for those records which were guidelines appropriate as determined by the reviewers.

METHODS

Outpatient chart notes were obtained from a 540 MB data set which includes records from June, 1994 through December, 1995 in the Lifetime Clinical Record™ (SMS, Malvern, PA) at Oregon Health Sciences University. We confined the records to clinics that could reasonably be expected to make diagnostic and therapeutic decisions regarding the patients' asthma: Family Practice, Pediatrics, General Internal Medicine, Allergy, and Pulmonary. A rules-based PERL partitioning program was constructed to identify records containing the strings "asthma" or "reactive airway(s) disease." Since asthma was not necessarily the primary reason for the patient encounter, further rules were constructed in a manual, iterative fashion to refine the selection process to eliminate unwanted records. For example, records which only contained strings such as "family history of asthma" or "1. Asthma. In remission." were eliminated. A 50-record training set and a 50-record test set were built by applying the partitioning program to the records database. No alterations, such as spelling corrections, were made to the training or test sets.

The training and test sets were then submitted for review to three pulmonary specialists (two physicians and one pulmonary nurse practitioner). A copy of the NHLBI/NAEP guidelines for the diagnosis of asthma as they appear in TABLE 1 were provided to the

reviewers. The reviewers were asked to review the records using the guidelines and their clinical

judgement to determine: 1) if the guidelines applied to the record, 2) the asthma severity based on information in the record, and 3) the need for inhaled anti-inflammatory agents. Majority provider opinion was used as the scoring reference standard. This required at least two of the three reviewers having identical responses to the question. If all three reviewers differed in their responses, the answer was coded as "can't tell." If the record was not appropriate for the guideline, questions 2 and 3 were coded as "not applicable."

A scoring algorithm constructed in PERL contained terms that related to the disease entity, patient history, symptoms, signs, and tests which were found in asthma rating scales including the NHLBI/NAEP guidelines. Medications used to treat asthma were also included in the scoring algorithm. Each occurrence of the primary terms and synonyms was counted. Possible modifiers of primary terms were identified using the grep function in Unix on a set of 100 records relating to asthma which were not part of the training or test sets. Modifiers could occur either before or after the primary term. Multiple modifiers were identified for some terms and scoring hierarchies were developed. In general, modifiers indicating absence of the term such as "no" in the string "no expiratory wheezes" were given the highest precedence. Scoring each term was accomplished by manually varying the initial weights of the primary term and by the way in which the initial weights of the primary terms were changed by the hierarchical rules of the modifiers. This process was continued iteratively until the scoring algorithm was able to replicate the reviewers' judgements on need for inhaled anti-inflammatory agents for each of the records in the training set for which the reviewers felt the NHLBI/NAEP guidelines were appropriate. An example of a term weight and the way in which the weight is changed by the hierarchical rules of the modifiers is shown in Table 2 for the term "breath sounds."

TABLE 2.

Scoring Algorithm for Primary Term "breath sounds"

Initial score = 1

If strings "normal," "good," or "clear" occur BEFORE term, score = 0, no further scoring.

If string "distant" occurs BEFORE term, score = 3, no further scoring.

If strings "decreased," or "diminished" occur BEFORE term, score = 2, but if string "slightly" occurs BEFORE term, score = 1, or if string "markedly" occurs BEFORE term, score = 3.

The total count of terms that appeared in each record and the total weighted score for each record were recorded. Records that were determined to show that the patient should be on inhaled anti-inflammatory agents had a minimum count of 6 and had a minimum score/count ratio of 1.5. The score/count ratio was used to adjust for the length of the outpatient record, since more findings were likely to occur in longer records.

RESULTS

TABLE 3 summarizes the NHLBI/NAEP guidelines appropriateness ratings of the reviewers. A majority of records were guidelines-appropriate in both the training and test sets. Of the appropriate records, a large majority were believed to need inhaled anti-inflammatory agents.

TABLE 3.

Reviewer Ratings of Training and Test Sets for Guidelines Appropriateness

Training Set

Appropriate	Can't Tell	Not Applicable
29(58%)	13(26%)	8(16%)

Findings of Appropriate records (n=29)

Anti-inflammatory agent requirements identified
21(72%)
Need anti-inflammatory agent 18(62%)
Don't need anti-inflammatory agent 3(8%)
Can't tell need for anti-inflammatory agents 8(28%)

Test Set

Appropriate	Can't Tell	Not Applicable
26(52%)	15(30%)	9(18%)

Findings of Appropriate records (n=26)

Anti-inflammatory agent requirements identified
21(81%)
Need anti-inflammatory agent 19(73%)
Don't need anti-inflammatory agent 2(8%)
Can't tell need for anti-inflammatory agents 5(19%)

Kappa scores of interrater reliability on need for inhaled anti-inflammatory agents were calculated from the 55 records of the combined training and test sets which consensus indicated were appropriate for the guidelines for the judgement. Pairwise kappa scores varied between 0.11 and 0.30. The expert reviewers had significantly more difficulty on agreement as to the severity of asthma even when requirements for anti-inflammatory agents were agreed upon. The primary disagreements were related to the inability to discriminate differences between moderate and severe asthma as defined by the guidelines. Of the records in

both the training and tests sets where there was consensus on requirements for anti-inflammatory agents only 12 of 21 (57%) had consensus on asthma severity for each set. We did not try to match the ratings of asthma severity because of this variability.

The scoring algorithm correctly matched the reviewers judgements on need (or lack of need) for inhaled anti-inflammatory agents in 16 of 21 records (76%) from the test set for which anti-inflammatory need could be determined. Four of the five errors in the scoring algorithm failed to predict the need for inhaled anti-inflammatory agents and one erroneously predicted a need for inhaled anti-inflammatory agents which the reviewers did not feel was indicated. The test set included 19 records where the reviewers felt that inhaled anti-inflammatory agents were indicated. Eight of these patients were on inhaled anti-inflammatory agents. Our scoring algorithm correctly identified 7 of the 11 records (64%) as needing inhaled anti-inflammatory agents when the patient was not taking them.

The count of non-medication terms in guidelines-appropriate records (combined from training and test sets) averaged 9.06 (median = 8) and averaged 5.64 (median = 6) for those records where it was not clear if the guidelines should be applied. These values were significantly different ($t = 2.34$, $p = 0.02$). The count of medication terms averaged 3.60 (median = 3) for the guidelines-appropriate records and average 1.70 (median = 1) for the records that were not clearly appropriate for the guidelines, and these values were also significantly different ($t = 3.21$, $p = 0.002$). TABLE 4 shows the mean and median occurrence per record by category of each major term. Many terms occur infrequently such as peak flow or FEV₁.

TABLE 4.

Occurrence of Terms by Category in Combined Training and Test Sets (Records Appropriate for Guidelines or Can't tell) (n=82)

Non-Medication Categories		
Median/Record	Mean/Record	
Asthma/Reactive Airway(s) Disease	2.71	2
Wheeze	1.27	1
Cough	0.98	0
Dyspnea/Shortness of Breath	0.57	0
Peak Flow/FEV ₁	0.35	0
Air Movement	0.22	0
Emergency Room	0.21	0
Rales/Crackles	0.21	0
Sputum	0.17	0
Cyanosis	0.13	0
Rhonchi	0.12	0

TABLE 4. Continued

Methacholine/Mecholyl	0.11	0
Breath Sounds	0.09	0
Retractions	0.04	0

Medication Categories

Median/Record	Mean/Record	
Bronchodilators	1.28	1
Oral Corticosteroids	0.84	0
Inhaled Anti-inflammatory Agents	0.67	0
Theophyllines	0.17	0

DISCUSSION

For many outpatient records there is adequate information to determine if the NHLBI/NAEP guidelines are appropriate to be applied and to determine whether a patient should or should not be on inhaled anti-inflammatory agents based on a single encounter. This is important as 11 of 19 patients (58%) who the expert reviewers felt should be on inhaled anti-inflammatory agents in the test set were not on these agents. Our scoring algorithm correctly identified 7 of 11 of these patients not on inhaled anti-inflammatory agents. It would be expected that even a higher percentage of patients would inappropriately not be on inhaled anti-inflammatory agents if our training and test sets had been confined to notes from only primary care clinics.

Outpatient records that are not clearly guidelines appropriate have fewer terms associated with asthma for both medication and non-medication terms. Expert reviewers may rely on the inclusion of medication information to aid in their assessment of the appropriateness of applying the guidelines to the records. Therapeutic regimens have been used as a means of determining asthma severity.⁹ Those records in our training and test sets with more frequent medication terms were more likely to be judged by the expert reviewers as appropriate for the guidelines to be applied.

Many terms occur surprisingly infrequently. The NHLBI/NAEP guidelines use the frequency of symptoms, activity tolerance, nocturnal symptoms, school/work attendance, and measures of lung function to determine the severity of asthma. We found that much of this information is not recorded in outpatient records. It was likely that our expert reviewers relied on other information such as presence or absence of wheezing and medications used to determine the severity of asthma and need for inhaled anti-inflammatory agents. This finding may have relevance for guideline authors who may be basing their

guidelines on information health care providers do not collect or routinely use.

This study was limited by the size of the training and the test sets which were further reduced in size by the number of records which were not appropriate for or unclear whether the NHLBI/NAEP guidelines should be applied. The primary reasons for records being judged as not appropriate for the guidelines were related to the smoking history and smoking status of the patient. The reviewers felt that smoking patients were felt to possibly have some component of chronic obstructive pulmonary disease, and appropriateness of application of the asthma guideline was unclear. The reviewers also felt that guideline application was in doubt when asthma and reactive airway(s) disease were qualified by terms such as "possible" or "presumed" or "suspected." These modifications will be added to the initial partitioning tool to improve the selection process and improve the chances that a selected record will be appropriate for the NAEP guidelines. Larger data sets, data from other institutions, and increased numbers of expert reviewers will allow us to improve the partitioning and scoring algorithms.

It was hypothesized that a relatively small number of primary terms and their modifiers would contain the majority of information expert reviewers would need to evaluate asthma severity and need for inhaled anti-inflammatory medications. Clearly, some of the outpatient records do not include enough information to make these judgements. However, the reviewers were able to make the determinations based on a single encounter note a significant proportion of the time. We hope to apply statistical methods such as factor analysis to the terms in subsequent studies to evaluate the relative contribution of the terms and the association of the terms to each other. Simplification and improvement in weightings of the terms and their modifiers may be possible with these techniques.

CONCLUSION

This pilot project demonstrated that the application of clinical practice guidelines for the diagnosis and management of asthma to single encounter outpatient records is possible using a relatively simple approach to natural language processing. We were able to identify a relatively small number of primary terms, their modifiers, and then develop a scoring mechanism to approximate expert judgements on need for inhaled anti-inflammatory agents. Modest success was also demonstrated in being able to identify outpatient records that are appropriate for the application of the

NHLBI/NAEP guidelines for the diagnosis and treatment of asthma. Further development of these tools may allow for rapid screening of large outpatient textual medical record databases. This would allow for evaluation of practitioner practice patterns, identify patients who are not adequately treated, enable assessment of the impact of clinical practice guidelines, and allow for the comparison of various methods of introduction of clinical practice guidelines.

Acknowledgments

This work was supported by the PHS NIH National Library of Medicine Training Program in Health Informatics LM-07088 (Dr. Ertle) and by the PHS NIH National Library of Medicine Grant LM-05879 (Dr. Hersh).

References

1. Field MJ, Lohr, KN eds. *Clinical Practice Guidelines: Directions for a New Program*. Washington, D.C. National Academy Press; 1990.
2. Gergen PJ, Mullally DI, Evans R3: National survey of prevalence of asthma among children in the United States, 1976 to 1980. *Pediatrics* 1988;81:1-7.
3. Weitzman M, Gortmaker SL, Sobol AM, Perrin JM: Recent trends in the prevalence and severity of childhood asthma. *JAMA* 1992;268:2673-2677.
4. Kivity S, Shochat Z, Bressler R, Wiener M, Lerman Y: The characteristics of bronchial asthma among a young adult population. *Chest* 1995;108:24-27.
5. Sheffer AL, Taggart VS: The National Asthma Education Program. Expert Panel Report Guidelines for Diagnosis and Management of Asthma. National Heart, Lung, and Blood Institute. *Med Care* 1993;31:MS20-MS28.
6. Armstrong D, Fry J, Armstrong P: General Practitioner's Views of Clinical Guidelines for Management of Asthma. *Int J Quality in Health Care* 1994;6:199-202.
7. Van Der Windt DAWM, Nagelkerke AF, Bouter LM, Dankert-Roelse JE, Veerman AJP: Clinical Scores for acute asthma in pre-school children. A review of the literature. *J Clin Epi* 1994;47:635-646.
8. Rosier MJ, Bishop J, Nolan T, Robertson CF, Carlin JB, Phelan PD: Measurement of Functional Severity of Asthma in Children. *Am J Respir Crit Care Med* 1994;149:1434-1441.
9. Busse W, Maisiak R, Young KRJ: Treatment Regimen and Side Effects of Treatment Measures. *Am J Respir Crit Care Med* 1994;149:S44-S50.
10. Bishop J, Carlin J, Nolan T: Evaluation of the Properties and Reliability of a Clinical Severity Scale for Acute Asthma in Children. *J Clin Epi* 1992;45:71-76.
11. Redier H, Daures JP, Michel C, et al: Assessment of the Severity of Asthma by an Expert System.. *Am J Respir Crit Care Med* 1995;151:345-352.
12. Sager N, Lyman M, Bucknall C, N.Nhan H, Tick LJ: Natural Language Processing and the Representation of Clinical Data. *JAMIA* 1994;1:142-160.
13. Aronow DB, Cooley JR, Soderland S: Automated Identification of Episodes of Asthma Exacerbation for Quality Measurement in a Computer-Based Medical Record. *Proceedings of 19th Symposium on Computer Applications in Medical Care* 1995;1:309-313.